Activity 2
The Meaning of Genetic Variation

Focus: Students investigate variation in the *beta globin* gene by identifying base changes that do and do not alter function, and by using several print resources to consider the significance in different environments of the base change associated with sickle cell disease.

Major Concepts: The ultimate source of genetic variation is differences in DNA sequences. Most of those genetic differences do not affect how individuals function. Some genetic variation, however, is associated with disease, and some improves the ability of the species to survive changes in the environment. Genetic variation, therefore, is the basis for evolution by natural selection.

Objectives: After completing this activity, students will

- recognize that the extent of molecular variation between two people is only about 0.1 percent, but because of the large size of the human genome, this translates to about 3 million base differences;
- understand that most human genetic variation does not affect function;
- be able to explain that some human genetic variation is related to disease and provide an example; and
- be able to describe a benefit of human genetic variation and relate this benefit to human evolution by natural selection.

Prerequisite Knowledge: Students should understand basic Mendelian patterns of inheritance, especially autosomal recessive inheritance; the basic structure of DNA; the transcription of DNA to messenger RNA; and the translation of messenger RNA to protein.

Basic Science-Health Connection: Although the idea is made explicit only in annotations to teachers, this activity illustrates how advances in science and technology have allowed us to establish relationships between some genetic variations and particular phenotypes. For example, our understanding of the relationship between DNA and protein has allowed us to establish a relationship between a change in a single base pair and the symptomology of sickle cell disease. Similarly, our understanding of the basic biochemical mechanisms underlying the symptoms associated with sickle cell disease has provided important clues about possible strategies for clinical intervention. You may wish to make some of these points with your students as they complete the activity.

As discussed in *Understanding Human Genetic Variation*, there is considerable variation between the genomes of any two individuals, but only a small amount of that variation has any significant biological impact, that is, produces differences in function. The Human Genome Project will continue to illuminate

the extent of human genetic variation as well as the variations that have biological significance.

This activity uses an examination of variation in a 1,691-base segment of the *beta globin* gene to help students consider the extent of human genetic variation at the molecular level and the relationships between genetic variation and disease and between genetic variation and evolution.

Materials and Preparation

You will need to prepare the following materials before conducting this activity*:

- Master 2.1, *How Much Variation? Beta Globin Gene—Person A* (make 1 copy per student)
- Master 2.2, *How Much Variation? Beta Globin Gene—Person B* (make 1 copy per student)
- Master 2.3, *How Much Variation? Doing the Math* (make 1 copy per student)
- Master 2.4, *Exploring Sickle Cell Disease* (make 1 copy per student)
- Master 2.5, *Reference Database* (make 1 copy per student)
- Master 2.6, *Results of the Lindsey Test* (make 1 copy per team)

*Day 1, Step 12 describes an optional laboratory exercise that you may wish to conduct to enrich your students' understanding of molecular variation and the methods by which it can be identified and studied. Information about the materials required is provided on page 6.

Procedure

DAY 1

1. **Introduce the activity by asking students to identify the ultimate source of the variation they investigated in Activity 1.**

   Students should recognize that the ultimate source of genetic variation is differences in DNA sequences.

2. **Explain that in this activity, students will investigate human genetic variation at a molecular level and examine the impact of that variation on biological function. Distribute one copy of Masters 2.1, *How Much Variation? Beta Globin Gene—Person A*, and 2.2, *How Much Variation? Beta Globin Gene—Person B*, to each student. Explain that the sequences on these pages come from the *beta globin* gene in two different people.**

   Hemoglobin, the oxygen carrier in blood, is composed of four polypeptide chains, two alpha polypeptide chains and two beta polypeptide chains. The *beta globin* gene encodes the amino acid sequence for the beta chain. *Person A* and *Person B* each show 1,691 nucleotides from the "sense" strand of the gene (that is, the strand that does not serve as the template for transcription and thus has the same base sequence as the messenger RNA, with Ts substituted for Us). Both the sense strand of DNA and the messenger RNA are complementary to the DNA strand that serves as the template for transcription. We recommend that you

2

remind students that DNA is double-stranded, even though only one strand is shown here. Explain that geneticists use "short cuts" like this because, given the sequence of one DNA strand, they can infer the sequence of the complementary strand.

The *beta globin* gene is one of the smallest human genes that encodes a protein; the entire gene has only about 1,700 nucleotide pairs and includes just two introns. The sequences on *Beta Globin Gene—Person A* and *Beta Globin Gene—Person B* do not show the gene's promoter regions, but begin with the first sequences that are translated.

3. **Ask the students to read the information at the top of each page and then estimate the total number of bases on each page. Direct students to write their estimate at the top of the page.**

   The total number of bases on *each* page is 1,691. Students will need this number to complete their calculations in Step 6.

4. **Remind the students that the sequences on the masters come from the *beta globin* gene in two different people. Ask the students what they notice when they compare the sequence from person A with the sequence from person B.**

   Students should notice that most of the sequence appears to be exactly the same in both people. They also should notice that the bases that are in bold are different. If necessary, point out that these bases are at the same positions in each gene (that is, be sure that students realize that only these two bases, located in these specific positions, are different in the sequences from person A and person B).

5. **Point out that this sequence is only 1,691 bases long and the complete human genome is approximately 3 billion bases long. Ask the students how they might use the sequences for person A and person B and the total size of the human genome to estimate the extent of variation (the number of bases that differ) between person A and person B. Ask as well what assumption they would be making as they arrived at their estimate.**

   Students could estimate the extent of variation across the entire genome by calculating the percentage of difference between the two sequences shown for person A and person B, and then multiplying this percentage by 3 billion (the approximate number of bases in the human genome). This estimate assumes that the sequence shown displays a typical amount of variation.

6. **Distribute one copy of Master 2.3, *How Much Variation? Doing the Math,* to each student and direct the students to use the master as a guide to estimate this value.**

   If your students need help completing this estimate, suggest that they first try the example at the bottom of the master.

The proportion of sequence difference between person A and person B is $2/1{,}691 = .001$ (rounded off). To make this more concrete for your students, note that this means that about 1 base in every 1,000 is different. The percentage difference is $.001 \times 100 = 0.1$ percent.

The total number of base differences would be $3{,}000{,}000{,}000 \times .001 = 3{,}000{,}000$ or, in scientific notation, $3 \times 10^9 \times 10^{-3} = 3 \times 10^6$. That is, we could expect to find 3 million base differences in DNA sequence between any two people.

Note that the actual number of base differences between two people likely is somewhat higher than this because this estimate, based as it is on the approximate size of the human *genome* (one copy of each of the autosomes, plus the X, Y, and mitochondrial chromosome), does not take into consideration the fact that humans are diploid.

7. **Ask students what their estimates indicate about the extent of human genetic variation at the molecular level.**

Students should recognize that at the molecular level, humans are far more alike (about 99.9 percent of the bases are the same) than they are different (only about 0.1 percent of the bases are different). Students also should realize, however, that even a small percentage difference can represent a very large actual number of differences in something as large as the human genome.

If students have difficulty reaching these conclusions, help them by asking questions such as, "Based on this comparison, do you think that at the molecular level, people are more alike than they are different or vice versa?" and "How can a difference of only 0.1 percent (1 in 1,000) result in such a large number of differences (3 million differences)?"

8. **Explain that the rest of the activity focuses on this 0.1 percent difference between people. Ask students questions such as "Do you think these differences matter? What effect do you think they have? What might affect how much a specific difference matters?"**

These questions focus students' attention on the *significance* of the differences, instead of the *number* of differences. Remind students of the differences among people that they observed in Activity 1 and point out that most of these differences have their basis in a difference in the DNA sequence of particular genes (probably pierced versus non pierced body parts do not). To help them understand the magnitude of the number of differences between their DNA and that of another person, ask students if they think there are 3 *million* differences in appearance and biological functions between themselves and the person sitting next to them.

9. **Explain that studying the *beta globin* gene more closely will help students begin to answer these questions for themselves. Direct students to examine**

the sequences on *Beta Globin Gene—Person A* and *Beta Globin Gene—Person B* **again. Explain that the regions that show bases grouped in triplets are from the coding regions ("exons") of the gene, while the other regions are from the noncoding regions ("introns"). Then ask students which of the two base differences in bold is most likely to matter, and why.**

Most eukaryotic genes are composed of both coding and noncoding regions, which are transcribed into an initial messenger RNA. The noncoding introns are then spliced out of the RNA; other processing steps ultimately result in the mature messenger RNA that is translated into protein. Students should realize that the second base difference occurs in a noncoding region of the gene and is unlikely to have an impact on individuals. The first difference occurs in a coding region and is more likely to matter.

10. **Explain that although 3 million base differences sounds like a lot, most of these differences have no significant impact on individuals, either because they occur in a noncoding region or for another reason. Point out that most of these 3 million differences can only be detected by examining the DNA sequence.**

A major concept that students should understand from Day 1 of the activity is that most genetic differences do not affect how individuals function.

Students should now understand that while some base differences occur in coding regions and may result in an altered amino acid sequence in the protein coded for by a gene, others occur in noncoding regions where they likely have no impact. Point out that only a small percentage of the DNA sequences in the human genome are coding sequences. Furthermore, only a small percentage of the noncoding DNA sequences are regulatory sequences such as promoters or enhancers that can influence the amount of gene product that results from a given gene. The remaining DNA sequences (the majority of the total DNA sequences in the genome) have no known function. Most of the variations in DNA sequence occur in these latter sequences and have no detectable impact.*

If you wish to offer your students a more sophisticated understanding of why most DNA sequence differences have no impact, extend the discussion to include the following ideas. Even many of the differences that occur in coding regions have no impact. Only those differences that result in a change in amino acid sequence in a critical region of the protein (one that affects the function of the protein), or that result in a premature stop codon in the RNA (and thus a truncated protein) have a significant impact on the individual carrying that variation. As students will see in Day 2, those few differences that do affect individuals often have devastating consequences.

*You may wish to clarify for students the reason that most molecular variation occurs in noncoding regions. It is true that there are more noncoding than coding regions. However, the fundamental biological reason for the increased variability of noncoding regions is that there is

no selective pressure exerted on changes in these silent/nonfunctional regions. You also may wish to point out that some differences that occur in noncoding regions do have an impact. For example, several mutations within introns in the *beta globin* gene cause incorrect splicing of the messenger RNA and, as a result, several codons may be inserted into or omitted from the sequence, leading to nonfunctional beta globin polypeptides.

11. **Point out the codon in which the first difference between the two sequences occurs and tell students that person A has normal hemoglobin, while Person B has an abnormal hemoglobin that is associated with sickle cell disease. Explain that the single base difference in this codon determines whether a person has normal hemoglobin or sickle hemoglobin.**

If you wish to take the time, ask students to identify the actual amino acid difference between these two types of hemoglobin, based on the difference in the DNA sequence of the codon you identified. This is an opportunity for students to review the translation process and the genetic code. Remind them that the sequence they have is the same as the messenger RNA sequence, except it has Ts where the RNA would have Us. Normal hemoglobin has glutamic acid (RNA codon GAG) in the position where sickle cell hemoglobin has valine (RNA codon GUG).

12. **Tell students that in the next part of the activity, they will consider the consequences of the genetic variation that results in sickle cell disease. Distribute one copy each of Masters 2.4, *Exploring Sickle Cell Disease*, and 2.5, *Reference Database*, and direct students to organize into their teams and begin working on the questions.**

Give the students about 30 minutes to complete their research. Notice that most of the information they need is located in the *Reference Database* handout. If the students' textbooks have an adequate description of sickle cell disease, you may wish to have students use this resource instead of the *Reference Database*.

When students reach Question 2 on *Exploring Sickle Cell Disease*, they should explain how they intend to test the Lindsey twins. Give the team a copy of the test results (Master 2.6, *Results of the Lindsey Test*) after the students correctly explain the test they would have conducted.

Instead of giving students the results of the test they propose in Question 2 on *Exploring Sickle Cell Disease*, you may want them to complete the relevant laboratory themselves. Kits such as Ward's *Identification of Genetic Diseases* laboratory activity (catalog # 36WS374) are available that you can adapt to this purpose. If you plan to have your students complete the lab, schedule an additional one-half to one class period for the activity.

DAY 2

1. **Open the second half of the activity by directing students to meet in their teams to complete or review their answers to the questions on *Exploring Sickle Cell Disease*. After they have completed their work, convene a class discussion in which you invite students to share their answers to the questions.**

    Question 1a What are the primary symptoms of sickle cell disease? What happens in a person's body to cause these symptoms?

    People with sickle cell disease periodically experience symptoms that include severe pain and fever. The symptoms occur when the sickle hemoglobin (Hb S) inside red blood cells forms long crystals under conditions of low oxygen concentration. The red blood cells elongate and assume a "sickle" shape. The crystallized hemoglobin damages the cell membranes, causing them to burst easily. The misshapen cells also clog blood vessels. The result is the destruction of many red blood cells within a few hours and a disruption of oxygen transport that can lead to death.

    Question 1b How is Hb S (sickle hemoglobin) different from Hb A (normal hemoglobin)?

    Sickle hemoglobin (Hb S) has the amino acid valine in the position where normal hemoglobin (Hb A) has the amino acid glutamic acid.

    Question 1c How can this difference in hemoglobin be detected in the laboratory?

    Because of the difference in the amino acid sequence of Hb A and Hb S, the two forms of hemoglobin have different charges. The two forms can be separated using electrophoresis because Hb S moves more slowly in an electric field than Hb A.

    Question 1d What does this difference in hemoglobin tell you about the DNA of people whose cells make Hb S as compared with people whose cells make normal hemoglobin?

    The sequence of DNA that codes for hemoglobin in people whose cells make Hb S must be different from the sequence of DNA that codes for hemoglobin in people whose cells make Hb A. The allele that codes for Hb A has the nucleotide A at a place where the allele that codes for Hb S has the nucleotide T.

    Question 1e What is the difference between sickle cell disease and sickle cell trait? Demonstrate in your answer that you understand how sickle cell disease is inherited.

    People who have sickle cell disease have inherited two alleles for sickle cell hemoglobin, one from each of their parents. They are homozygous for the sickle cell hemoglobin allele. People who have

7

sickle cell trait have inherited one allele for sickle cell hemoglobin from one parent and one allele for normal hemoglobin from the other parent. They are heterozygous and usually have no symptoms.

Question 2 Use what you learned about sickle cell disease and trait to propose a way to determine whether Ms. Lindsey's twins have sickle cell trait. Explain your procedure to your teacher, then use the information provided on the handout your teacher will give you to determine the results of the test.

Students should explain the following procedure: Collect hemoglobin from Jason and from Sondra and determine the form(s) of hemoglobin each has using gel electrophoresis. "Standards," or controls, of normal and sickle hemoglobin should be included for comparison.

If a twin is normal, his or her hemoglobin will migrate on the gel in parallel with the Hb A standard. If a twin has sickle cell disease, his or her hemoglobin will migrate like the Hb S standard. If a twin is heterozygous (has sickle cell trait), his or her hemoglobin will contain two forms of hemoglobin, one that migrates like Hb A and one that migrates like Hb S.

Question 3 Write the dialogue for a brief (2–3 minute) scene in which you explain to Ms. Lindsey the results of the tests you ran on the twins, what these results say about the inheritance of the sickle cell trait in her family, and the implications of your findings for the twins' health.

Responses will vary, but students should indicate that Sondra has sickle cell trait and Jason has sickle cell disease. These results indicate that both Ms. Lindsey and her late husband also have/had sickle cell trait; that is, they are both heterozygous for the sickle cell allele and the normal allele, because neither of them are/were ill, but each of them must have given a sickle cell allele to Jason. Sondra should be fine, but Jason has sickle cell disease.

2. **Ask students what their study of the *beta globin* gene and sickle cell disease has illustrated about human genetic variation.**

Students should be able to describe the extent of genetic variation from one person to another and should be able to explain that most of these differences do not have a significant biological impact. Students should recognize, however, that some variation (for example, the single base change associated with sickle cell disease) produces negative consequences.

3. **Summarize the students' answers by saying, "So you are saying that most variation does not make a difference and that some variation is negative. Is it possible that some variation also is positive?" Entertain several answers to this question.**

Highlight the contribution of basic science to the improvement of personal and public health by asking your students whether an early-20th-century physician would have answered Ms. Lindsey's question in this manner. The answer, of course, is no. The first observation of sickle-shaped cells was made in 1910, but the molecular basis of the disease was not worked out until 1949. You may also note that direct diagnosis of this disease through DNA analysis of a person's genotype was made possible in the mid-1980s.

Collect the students' written scenes or have each team perform its scene for the class as a way to assess students' understanding of the inheritance of sickle cell disease.

Most students will recognize that it is possible that some variation is positive.

4. **Ask students, as a final challenge, to imagine that they are doctors practicing in Cameroon, in west-central Africa. Direct them to return to the** *Reference Database* **to compare the incidence of sickle cell disease in Cameroon with its incidence in the United States and to determine how scientists explain the difference.**

The incidence of sickle cell disease among black Africans is as much as 16 times higher than the incidence among African Americans (4 percent compared with .25 percent). Scientists believe this difference is related to the occurrence of malaria in many parts of Africa. People who are homozygous for the normal allele for hemoglobin die of malaria more often than people who are heterozygous for the normal allele and the sickle allele for hemoglobin. Thus, more heterozygotes live than people who are homozygous for the normal allele, and they pass their allele for sickle hemoglobin on to many of their children. The result is that the proportion of this allele is higher in the population than it would be if there were no threat of malaria. In contrast, in the United States where there is practically no threat of malaria, people with sickle cell trait (heterozygotes) are no more likely to live than those who are homozygous for the normal allele for hemoglobin. So the proportion of the allele for sickle hemoglobin remains at a very low level in the population because those individuals who inherit two copies of this allele suffer with sickle cell disease and frequently die before passing their alleles on to any offspring.

5. **Ask students how this information would change what they would say to Ms. Lindsey.**

The only thing that would change is the implication of the findings for the twins' health. Jason will still have sickle cell disease, but Sondra should have enhanced resistance to malaria.

6. **Close the activity by challenging the class to answer the following questions:**

• **Will natural selection favor the survival of people who produce Hb S or people who produce normal hemoglobin?**

The critical variable here is the environment in which the Hb S variation is expressed. In environments where malaria is endemic, those who are heterozygous for the Hb S allele (HbA/Hb S) will be more resistant to malaria than are those who are homozygous for the Hb A allele (Hb A/Hb A). Evidence indicates that natural selection has favored the heterozygous state in those environments, therefore maintaining the Hb S allele in relatively high frequencies in some populations during the course of human evolution. In a nonmalarial environment, there is no known selective advantage to carrying the

This challenge will reinforce a major concept of this activity. Some genetic variation has negative consequences for individuals. However, some genetic variation improves the ability of the species to survive changes in the environment. Genetic variation is the basis for evolution by natural selection.

Hb S allele in the heterozygous state. Those who are homozygous for the Hb S allele (Hb S/Hb S) likely will experience the symptoms of sickle cell disease in any environment.

- **All populations have genetic variations that lead to increased incidence of particular disorders (for example, cystic fibrosis among Caucasians of European ancestry, Tay-Sachs disease among Eastern-European Jews, and a particular type of thalassemia—a blood disorder—among Asians). Challenge students to explain why such apparently harmful variations have been maintained in those populations.**

Although most genetic variation is meaningless, some of it is harmful and some of it is beneficial because it improves the ability of the species to survive changes in the environment. The most likely explanation for these examples is the one that has been most clearly established for sickle cell disease: There is a survival/reproductive advantage for people who are heterozygous as compared with those who are homozygous for the normal allele.

In the case of cystic fibrosis, there is good evidence that those who carry one CFTR allele associated with the disease have increased resistance to typhus, a common killer in Europe in past centuries. There also is circumstantial evidence that those who have one allele associated with Tay-Sachs disease may be more resistant to tuberculosis than those who are homozygous for the normal allele.

How Much Variation?
Beta Globin Gene—Person A

This page contains the DNA base sequence for *part* of a gene called *beta globin*. Hemoglobin, the oxygen carrier in blood, is composed of four polypeptide chains, two alpha polypeptide chains, and two beta polypeptide chains. The *beta globin* gene encodes the amino acid sequence for the beta chain. The complete gene is about 1,700 DNA bases long.

Read the sequence from left to right across the page.

```
ATG GTG GAC CTG ACT CCT GAG GAG AAG TCT GCC GTT ACT GCC CTG TGG GGC AAG GTG
AAC GTG GAT GAA GGT GGT GTT GAG GCC CTG GGC AGGTTGGTATCAAGGTTACAAGACAGGTTTAAG
GAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATT
GGTCTATTTTCCCACCCTTAG G CTG CTG GTG GTC TAC CCT TGG ACC CAG AGG TTC TTT GAG
TCC TTT GGG GAT CTG TCC ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG AAG GCT
CAT GGC AAG AAA GTG CTC GGT GCC TTT AGT GAT GGC CTG GCT CAC CTG GAC AAC CTC
AAG GGC ACC TTT GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG CAC GTG GAT CCT
GAG AAC TTC AGG GTGAGTCTATGGGACGCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTC
ATAGGAAGGGGAGAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCTCA
GGATCGTTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCTTTTTTTTTCT
TCTCCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAG
TAACTTAAAAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTTGGAATATATGTGTGCTTATTTGCATATTCAT
AATCTCCCTACTTTATTTTCTTTTATTTTTAATTGATACATAATCATTATACATATTTATGGGTTAAAGTGTAATGTT
TTAATATGTGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATA
TACTTTTTTGTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGC
CTCTTTGCACCATTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATTTCTGCATATAAATATTT
CTGCATATAAATTGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTA
TTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATCT
TCCTCCCACAG CTC CTG GGC AAC GTG CTG GTC TGT GTG CTG GCC CAT CAC TTT GGC AAA
GAA TTC ATC CCA CCA GTG CAG GCT GCC TAT CAG AAA GTG GTG GCT GGT GTG GCT AAT
GCC CTG GCC CAC AAG TAT CAC TAA GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTT
CCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAACATTTA
TTTTCATTGCAATGATGTATTTAAATTATTTCTGAATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCATTTAAA
ACATAAAGAAATGATGAGCTGTTCAAACCTTGGGAAAATACACTATATCTTAAACTCCATGAAGAA
```

How Much Variation?
Beta Globin Gene—Person B

This page contains the DNA base sequence for *part* of a gene called *beta globin*. Hemoglobin, the oxygen carrier in blood, is composed of four polypeptide chains, two alpha polypeptide chains, and two beta polypeptide chains. The *beta globin* gene encodes the amino acid sequence for the beta chain. The complete gene is about 1,700 DNA bases long.

Read the sequence from left to right across the page.

```
ATG GTG GAC CTG ACT CCT G T G GAG AAG TCT GCC GTT ACT GCC CTG TGG GGC AAG GTG
AAC GTG GAT GAA GGT GGT GTT GAG GCC CTG GGC AGGTTGGTATCAAGGTTACAAGACAGGTTTAAG
GAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATT
GGTCTATTTTCCCACCCTTAG G CTG CTG GTG GTC TAC CCT TGG ACC CAG AGG TTC TTT GAG
TCC TTT GGG GAT CTG TCC ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG AAG GCT
CAT GGC AAG AAA GTG CTC GGT GCC TTT AGT GAT GGC CTG GCT CAC CTG GAC AAC CTC
AAG GGC ACC TTT GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG CAC GTG GAT CCT
GAG AAC TTC AGG GTGAGTCTATGGGAC C CTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTC
ATAGGAAGGGGAGAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAGTCTCA
GGATCGTTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCTTTTTTTTTCT
TCTCCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAG
TAACTTAAAAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTTGGAATATATGTGTGCTTATTTGCATATTCAT
AATCTCCCTACTTTATTTTCTTTTATTTTTAATTGATACATAATCATTATACATATTTATGGGTTAAAGTGTAATGTT
TTAATATGTGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATA
TACTTTTTTGTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGC
CTCTTTGCACCATTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATTTCTGCATATAAATATTT
CTGCATATAAATTGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTA
TTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATCT
TCCTCCCACAG CTC CTG GGC AAC GTG CTG GTC TGT GTG CTG GCC CAT CAC TTT GGC AAA
GAA TTC ATC CCA CCA GTG CAG GCT GCC TAT CAG AAA GTG GTG GCT GGT GTG GCT AAT
GCC CTG GCC CAC AAG TAT CAC TAA GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTT
CCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTA
TTTTCATTGCAATGATGTATTTAAATTATTTCTGAATATTTTACTAAAAAGGGAATGTGGGAGGTCAGTGCATTTAAA
ACATAAAGAAATGATGAGCTGTTCAAACCTTGGGAAAATACACTATATCTTAAACTCCATGAAAGAA
```

How Much Variation?
Doing the Math

Calculate the amount of variation in the DNA in the *beta globin* gene between person A and person B. If you need help, use the example below as a guide.

1. How many bases are different between the sequence shown for person A and the sequence shown for person B? _____

   How many total bases are in the sequence? _____ (Your teacher will give you this number.)

   Divide the number of different bases by the total number of bases in the sequence.

   $$\frac{\text{number of different bases}}{\text{total number of bases}} = \text{_____} = \text{_____}$$

2. The percentage difference is _____ × 100 = _____ %.

3. The human genome has about 3 billion bases. Assume that the degree of difference you just calculated applies across the entire genome. How many total base differences would you expect to find between person A and person B?

   3,000,000,000 × _____ = _____ total differences

   or, in scientific notation, $3 \times 10^9 \times$ _____ = _____

Example

The sophomore class at Roosevelt High School in Metropolitan City is one of five high schools that conduct two community service projects each year, one in the fall and one in the spring. This fall, 150 students from Roosevelt High signed up to help. The same number signed up in the spring, but 30 of the students were different. What percentage of the students was different between the fall group and the spring group?

1. To calculate the percentage difference, first divide the number of different students in the spring by the total number of students in the group:

   $$\frac{\text{number of different students}}{\text{total number of students}} = \frac{30}{150} = .2$$

2. Convert this result to a percentage by multiplying by 100: .2 × 100 = 20%

3. The sophomore classes at all five high schools combined include about 3,000 students. Assume that the degree of difference between the students who signed up for the community service projects in the fall and spring across all five high schools is the same as it was at Roosevelt High. How many different students would you expect to find in total between the fall and spring projects?

   3,000 × 20% = 600 different students

Exploring Sickle Cell Disease

Imagine that you are a family practice physician and that an African-American woman, Audrey Lindsey, and her family are your patients. Just before her twins, Sondra and Jason, were born, Ms. Lindsey's husband, also African-American, died in an automobile accident. His parents were physiologically normal, but he had a brother who died of sickle cell disease at the age of 19. Ms. Lindsey explains to you that it is important to her to know whether her twins carry the allele associated with sickle cell disease.

1. Use the *Reference Database* and any other resources that are available to you (for example, your textbook) to answer the following questions:

    a. What are the primary symptoms of sickle cell disease? What happens in a person's body to cause these symptoms?

    b. How is Hb S (sickle hemoglobin) different from Hb A (normal hemoglobin)?

    c. How can this difference in hemoglobin be detected in the laboratory?

    d. What does this difference in hemoglobin tell you about the DNA of people whose cells make Hb S as compared with people whose cells make normal hemoglobin?

    e. What is the difference between sickle cell disease and sickle cell trait? Demonstrate in your answer that you understand how sickle cell disease is inherited.

2. Use what you learned about sickle cell disease and trait to propose a way to determine whether Ms. Lindsey's twins have sickle cell trait. Explain your procedure to your teacher, then use the information provided on the handout your teacher will give you to determine the results of the test.

3. Write the dialogue for a brief (2–3 minute) scene in which you explain to Ms. Lindsey the results of the tests you ran on the twins, what these results say about the inheritance of the sickle cell trait in her family, and the implications of your findings for the twins' health.

Reference Database

Sickle Cell Disease—Definition

Sickle cell disease is a genetic disorder that affects approximately 1 out of every 625 African Americans in the United States. It is caused by a single amino-acid change in a protein called hemoglobin.

Hemoglobin is the major protein inside red blood cells. Its primary function is to transport oxygen. When the oxygen concentration in the blood decreases, the defective hemoglobin molecule forms long crystals inside the red blood cell. These crystals cause the red blood cells to elongate and assume a "sickle" shape. The crystallized hemoglobin also damages the cell membrane so that the cells become very fragile.

Sickle Cell Disease—Incidence (Africa)

In some parts of Africa, about 4 percent of black Africans have sickle cell disease.

Why does sickle cell disease occur more frequently among black Africans than among African Americans? Scientists believe that this difference is related to the threat of a fatal form of malaria that occurs in many parts of Africa. Studies reveal that people who are homozygous for the normal allele for hemoglobin (Hb A/Hb A) often die of malaria. However, people with sickle cell trait (people who are heterozygous, Hb A/Hb S), do not contract the fatal form of malaria. Thus, more heterozygotes live than do people who are homozygous for the normal allele, and these people often pass the sickle cell allele on to their children. This phenomenon keeps the incidence of the sickle cell allele in the population higher than it would be if there were no threat of malaria.

Sickle Cell Disease—Incidence (United States)

About one-quarter of 1 percent (0.25%) of African Americans are homozygous for the sickle cell hemo-globin allele and have sickle cell disease.

Why is the incidence of sickle cell disease in the United States so low? It is low because people with sickle cell disease often die in childhood or early adulthood, before they have had children. Thus, many people with sickle cell disease do not pass this allele on to children. Instead, most inheritance of the allele is from a parent who is heterozygous for the allele to one or more of his or her children.

Sickle Cell Disease—Inheritance

Sickle cell disease results when a person inherits an allele for sickle cell hemoglobin from each of his or her parents. This inheritance pattern means that the person is homozygous for sickle cell hemoglobin and that his or her body does not produce any normal hemoglobin, but only sickle cell hemoglobin.

Geneticists show the inheritance pattern of sickle cell disease by using symbols to represent the allele for normal hemoglobin (Hb A) and the allele for sickle cell hemoglobin (Hb S). A person with normal hemo-globin has inherited one allele for normal hemoglobin from each parent and so has the genotype Hb A/Hb A. In contrast, a person who has sickle cell disease has inherited one sickle cell allele from each parent and has the genotype Hb S/Hb S.

But what about a person who inherits an allele for normal hemoglobin from one parent and an allele for sickle cell hemoglobin from the other parent? This person has the genotype Hb A/Hb S and is said to

have *sickle cell trait*. Although some of the hemoglobin in this person's body is sickle cell hemoglobin, the rest of the hemoglobin is normal, and the person usually exhibits no symptoms of the disease.

Sickle Cell Disease—Structure of Altered Hemoglobin

The hemoglobin that is made in the bodies of people with sickle cell disease (Hb S) differs from normal hemoglobin (Hb A) in just one amino acid. In normal hemoglobin, this amino acid is a glutamic acid. In sickle cell hemoglobin, it is a valine.

Sickle Cell Disease—Symptoms

People with sickle cell disease may experience symptoms such as severe pain, fever, and even death. These symptoms occur when the misshapen red blood cells that form under conditions of low oxygen concentration clog blood vessels and burst.

Such patients frequently experience a vicious cycle of events called a "sickle cell disease crisis" in which low oxygen concentration causes sickling, which causes ruptured red blood cells, which in turn, causes even lower oxygen concentrations in the body and still more sickling and red blood cell destruction. This process leads to a serious loss of red blood cells within a few hours and can cause death.

Sickle Cell Disease—Testing

Normal hemoglobin (Hb A) and sickle cell hemoglobin (Hb S) differ in just one amino acid. Hb S has valine in the position where Hb A has glutamic acid. This difference results from a difference between the DNA sequence of the allele that codes for normal hemoglobin and the sequence of the allele that codes for sickle cell hemoglobin.

The difference in one amino acid between Hb A and Hb S causes a difference in the electrical charge of the two forms of hemoglobin. Hb A has a greater negative charge than Hb S. This difference can be used to distinguish Hb A and Hb S in a process called electrophoresis. Hemoglobin from individuals suspected of having sickle cell disease or sickle cell trait is placed on a gelatinous slab (gel) beside standards of Hb A and Hb S. An electrical charge is applied across the gel, and the proteins move through the gel toward the positive end of the electrical field at a rate based on their size and charge. Because Hb A has a greater negative charge than Hb S, it will move further through the gel. After electrophoresis, the gel is removed and stained with a solution that adheres to proteins, revealing "bands" of stain at the positions to which the hemoglobin has migrated.

Doctors diagnose sickle cell disease by comparing the banding position of hemoglobin from an individual with the banding positions of Hb A and Hb S standards. An individual who is homozygous for Hb A will have only one protein band on the gel, at the same position as the Hb A standard, whereas an individual who is homozygous for Hb S (and has sickle cell disease) will also have one protein band on the gel, but at the same position as the Hb S standard. A heterozygous individual will have two protein bands, one at each position.

Results of the Lindsey Test

Examine the following results to determine Sondra's and Jason's status with respect to sickle cell trait.
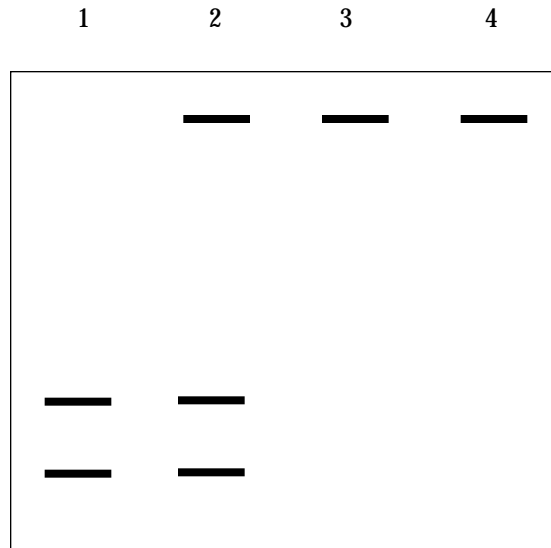
Lane

1   Standard—DNA from allele for Hb A

2   DNA from Sondra Lindsey

3   DNA from Jason Lindsey

4   Standard—DNA from allele for Hb S